The Next Era of Research Productivity Evaluation: Forming a Cohesive Metric

Ashlea Higgs

- Indexing: Provide terms for indexing the submission

  - Academic discipline and sub-disciplines: STI policy

  - Keywords: evaluating research productivity; performance metrics

  - Type, method or approach: theoretical modeling approach for multi-factorial metric

Abstract:

There is a clear need for performance measures related to research productivity in academic institutions and government funding and policymaking bodies worldwide today. While the general consensus on the need for performance measures is strong, however, there is a strong debate about what types of content and what kind of specific metrics should be included in such measures. What would a reasonable performance measure look like, and how would it be applied? To individual faculty members, to groups of scientists working together, or to entire departments or institutions? Many advocates of new metrics state that the metrics making up any productivity measure need to be more quantitative and objective. Some in this camp would argue that the traditional metrics, including reputation amongst peers, quality of publications, enthusiasm and effectiveness in mentoring younger scientists, ability to write and explain his or her research area clearly, and quality of collaborators, are too qualitative and subjective. Others would argue to the contrary that these supposedly subjective metrics capture critical qualities about a researcher's overall effectiveness, and that excessive reliance on quantitative metrics, such as bibliographic analysis and use of grants funding data, has the potential to be highly reductionist and unfair in its implementation.

We propose that any single performance metric is inadequate to capture the true dynamics of a researcher's productivity, but that a combinatorial approach between qualitative and quantitative metrics, with careful selection of each, can lead to a reasonable and standards-based approach to evaluating research productivity. We would like to speak on how such a metric could be created, what specific elements could be used, and how its efficacy could be measured during its initial implementation. This topic is certain to encourage passionate debate and multiple perspectives, and as such we feel it is a good starting point to engage the group assembled.

Quantitative Metrics

- Traditional bibliometric analysis (citation counts, impact factor, cites per article, etc.)
  - Document output is a common metric, and a discussion of metrics for individuals would be remiss without including the h-index. In a cohesive metric, should career length be a factor? Subject matter?
- Bibliographic-based co-citation analysis

    - In interviews, administrators appear to be more receptive to this metric than research management.
- Funding metrics: direct and/or indirect costs per square foot, funding per researcher, funding "density" metrics
    - Amount and success in securing funding is worthwhile for evaluation of research management. Costs per square foot and density metrics are more appropriate on the administrator level. True or false? Why?
- Survey results from students / post-docs ranking teaching and mentoring abilities of individual faculty members
    - Relevant for a more senior researcher
- Other scholarly output: patents, conference papers, books published, etc.
    - Importance of output varies by field. For example books in social sciences and arts & humanities, patents in applied sciences.
- Prizes: Field Medal, Nobel Prize, Turing Award, etc.
- Other distinctions: membership in National Academy of Sciences, etc.
- NIH / NSF funding rankings by institution
- Total research dollars / Total Federal research dollars (as per annual reports, NSF studies)
- International research university rankings (global: Times Higher Education, Shanghai Jiao Tong List; U.S. only: U.S. News and World Report)

Qualitative Metrics

- Departmental and peer review of an individual researcher's work
- Academic committees that review scholarly accomplishment and make recommendations on promotion and tenure
- Teaching capabilities as judged by peers
- Outcomes analysis of a researcher's success in training his or her post-docs
- University standing: students / faculty / employers / corporations partnering in research / government bodies
- Diversity of institution: social / ethnic / economic / ideological
- Degree of internationalization / interdisciplinary work

For consideration by the roundtable:

1. Which metrics provide the most versatile combinatorial metric across subject areas?
    a. As research becomes more interdisciplinary, does it become more difficult to reach consensus on a cohesive metric, or rather more necessary?
    b. Which factors of the ones mentioned above are the most determinative of research quality? Why? Is there a "golden ratio" of quantitative to qualitative factors? If so, are these factors broadly applicable or do they differ significantly by discipline?
2. What is more difficult to agree on – metrics for individuals or institutions? Why?
3. Some metrics may have more theoretical than practical value. Ability to implement is a factor which cannot be ignored. Which metrics require more political will? Which metrics have the greatest gap between current data availability and importance (i.e. very difficult to collect or without an existing, standardize body of data, yet very high importance)?